

SEP 23 2003

TECH CE



1600

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/762,261C

DATE: 09/16/2003

TIME: 16:45:03

Input Set : A:\h m jackson 5001 us.txt

Output Set: N:\CRF4\09162003\I762261C.raw

3 <110> APPLICANT: Quinnan, Gerald V.
4 Zhang, Peng Fei
6 <120> TITLE OF INVENTION: Expression and Characterization of HIV-1 Envelope
7 Protein Associated with a Broadly Reactive Neutralizing
8 Antibody Response
10 <130> FILE REFERENCE: 44508-5001-US
12 <140> CURRENT APPLICATION NUMBER: US 09/762,261C
C--> 13 <141> CURRENT FILING DATE: 2001-05-29
15 <150> PRIOR APPLICATION NUMBER: US 60/095,267
16 <151> PRIOR FILING DATE: 1998-08-04
18 <150> PRIOR APPLICATION NUMBER: PCT/US99/17596
19 <151> PRIOR FILING DATE: 1999-08-04
21 <160> NUMBER OF SEQ ID NOS: 23
23 <170> SOFTWARE: PatentIn Ver. 2.1
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 866
27 <212> TYPE: PRT
28 <213> ORGANISM: Human immunodeficiency virus type 1
30 <220> FEATURE:
31 <223> OTHER INFORMATION: R2 strain envelope protein (gp 160)
33 <400> SEQUENCE: 1
34 Met Arg Val Lys Gly Ile Arg Arg Asn Tyr Gln His Trp Trp Gly Trp
35 1 5 10 15
37 Gly Thr Met Leu Leu Gly Leu Leu Met Ile Cys Ser Ala Thr Glu Lys
38 20 25 30
40 Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr
41 35 40 45
43 Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Ala
44 50 55 60
46 His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
47 65 70 75 80
49 Gln Glu Val Glu Leu Val Asn Val Thr Glu Asn Phe Asn Met Trp Lys
50 85 90 95
52 Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp
53 100 105 110
55 Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
56 115 120 125
58 Asn Cys Thr Asp Leu Arg Asn Thr Thr Asn Thr Asn Asn Ser Thr Asp
59 130 135 140
61 Asn Asn Asn Ser Asn Ser Glu Gly Thr Ile Lys Gly Gly Glu Met Lys
62 145 150 155 160
64 Asn Cys Ser Phe Asn Ile Ala Thr Ser Ile Gly Asp Lys Met Gln Lys
65 165 170 175

ENTERED

RAW SEQUENCE LISTING

DATE: 09/16/2003

PATENT APPLICATION: US/09/762,261C

TIME: 16:45:03

Input Set : A:\h m jackson 5001 us.txt

Output Set: N:\CRF4\09162003\I762261C.raw

```

67 Glu Tyr Ala Leu Leu Tyr Lys Leu Asp Ile Glu Pro Ile Asp Asn Asp
68      180      185      190
70 Asn Thr Ser Tyr Arg Leu Ile Ser Cys Asn Thr Ser Val Ile Thr Gln
71      195      200      205
73 Ala Cys Pro Lys Ile Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala
74      210      215      220
76 Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys Lys Phe Ser Gly
77 225      230      235      240
79 Lys Gly Ser Cys Lys Asn Val Ser Thr Val Gln Cys Thr His Gly Ile
80      245      250      255
82 Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu
83      260      265      270
85 Glu Glu Val Val Ile Arg Ser Glu Asn Phe Thr Asn Asn Ala Lys Thr
86      275      280      285
88 Ile Ile Val Gln Leu Arg Glu Pro Val Lys Ile Asn Cys Ser Arg Pro
89      290      295      300
91 Asn Asn Asn Thr Arg Lys Ser Ile Pro Met Gly Pro Gly Arg Ala Phe
92 305      310      315      320
94 Tyr Thr Thr Gly Gln Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn
95      325      330      335
97 Ile Ser Lys Thr Asn Trp Thr Asn Ala Leu Lys Gln Val Val Glu Lys
98      340      345      350
100 Leu Gly Glu Gln Phe Asn Lys Thr Lys Ile Val Phe Thr Asn Ser Ser
101      355      360      365
103 Gly Gly Asp Pro Glu Ile Val Thr His Ser Phe Asn Cys Ala Gly Glu
104 370      375      380
106 Phe Phe Tyr Cys Asn Thr Thr Gln Leu Phe Asp Ser Ile Trp Asn Ser
107 385      390      395      400
109 Glu Asn Gly Thr Trp Asn Ile Thr Arg Gly Leu Asn Asn Thr Gly Arg
110      405      410      415
112 Asn Asp Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Arg
113      420      425      430
115 Trp Gln Glu Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Lys Gly Asn
116      435      440      445
118 Ile Ser Cys Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly
119      450      455      460
121 Gly Lys Asp Asp Asn Ser Arg Asp Gly Asn Glu Thr Phe Arg Pro Gly
122 465      470      475      480
124 Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys
125      485      490      495
127 Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg
128      500      505      510
130 Arg Val Val Gln Arg Glu Glu Arg Ala Val Gly Leu Gly Ala Met Phe
131      515      520      525
133 Phe Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Val
134      530      535      540
136 Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln
137 545      550      555      560
139 Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln

```

RAW SEQUENCE LISTING

DATE: 09/16/2003

PATENT APPLICATION: US/09/762,261C

TIME: 16:45:03

Input Set : A:\h m jackson 5001 us.txt

Output Set: N:\CRF4\09162003\I762261C.raw

```

140          565          570          575
142 Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Ile Leu Ala Val
143          580          585          590
145 Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser
146          595          600          605
148 Gly Lys Leu Ile Cys Thr Thr Thr Val Pro Trp Asn Ala Ser Trp Ser
149          610          615          620
151 Lys Asn Lys Thr Leu Glu Ala Ile Trp Asn Asn Met Thr Trp Met Gln
152 625          630          635          640
154 Trp Asp Lys Glu Ile Asp Asn Tyr Thr Ser Leu Ile Tyr Ser Leu Ile
155          645          650          655
157 Glu Glu Ser Pro Ile Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu
158          660          665          670
160 Leu Asp Lys Trp Ala Asn Leu Trp Asn Trp Phe Asp Ile Ser Asn Trp
161          675          680          685
163 Leu Trp Tyr Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Val Gly
164          690          695          700
166 Leu Arg Ile Val Phe Val Val Leu Ser Ile Val Asn Arg Val Arg Gln
167 705          710          715          720
169 Gly Tyr Ser Pro Leu Ser Phe Gln Thr Arg Leu Pro Ala Pro Arg Gly
170          725          730          735
172 Pro Asp Arg Pro Glu Glu Ile Glu Glu Gly Gly Asp Arg Asp Arg
173          740          745          750
175 Asp Arg Ser Gly Leu Leu Val Asp Gly Phe Leu Thr Leu Ile Trp Val
176          755          760          765
178 Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Leu
179          770          775          780
181 Leu Leu Ile Val Thr Arg Ile Val Glu Leu Leu Gly Arg Arg Gly Trp
182 785          790          795          800
184 Glu Ile Leu Lys Tyr Trp Trp Asn Leu Leu Gln Tyr Trp Ser Gln Glu
185          805          810          815
187 Leu Lys Asn Ser Ala Val Ser Leu Phe Asn Ala Thr Ala Ile Ala Val
188          820          825          830
190 Ala Glu Gly Thr Asp Arg Val Ile Gln Val Leu Gln Arg Val Gly Arg
191          835          840          845
193 Ala Leu Leu His Ile Pro Thr Arg Ile Arg Gln Gly Leu Glu Arg Ala
194          850          855          860
196 Leu Leu
197 865
200 <210> SEQ ID NO: 2
201 <211> LENGTH: 17
202 <212> TYPE: PRT
203 <213> ORGANISM: Human immunodeficiency virus type 1
205 <220> FEATURE:
206 <223> OTHER INFORMATION: segment of R2 strain V3 domain
208 <400> SEQUENCE: 2
209 Lys Ser Ile Pro Met Gly Pro Gly Arg Ala Phe Tyr Thr Thr Gly Gln
210 1          5          10          15
212 Ile

```

RAW SEQUENCE LISTING

DATE: 09/16/2003

PATENT APPLICATION: US/09/762,261C

TIME: 16:45:03

Input Set : A:\h m jackson 5001 us.txt

Output Set: N:\CRF4\09162003\I762261C.raw

```

216 <210> SEQ ID NO: 3
217 <211> LENGTH: 35
218 <212> TYPE: PRT
219 <213> ORGANISM: Human immunodeficiency virus type 1
221 <220> FEATURE:
222 <223> OTHER INFORMATION: R2 strain V3 domain
224 <400> SEQUENCE: 3
225 Cys Ser Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Pro Met Gly Pro
226   1           5           10           15
228 Gly Arg Ala Phe Tyr Thr Thr Gly Gln Ile Ile Gly Asp Ile Arg Gln
229           20           25           30
231 Ala His Cys
232           35
235 <210> SEQ ID NO: 4
236 <211> LENGTH: 35
237 <212> TYPE: PRT
238 <213> ORGANISM: Human immunodeficiency virus type 1
240 <220> FEATURE:
241 <223> OTHER INFORMATION: V3 domain of strain 93TH966.8
243 <400> SEQUENCE: 4
244 Cys Thr Arg Pro Ser Asn Asn Thr Arg Thr Ser Thr Thr Ile Gly Pro
245   1           5           10           15
247 Gly Gln Val Phe Tyr Arg Thr Gly Asp Ile Thr Gly Asn Ile Arg Lys
248           20           25           30
250 Ala Tyr Cys
251           35
-----
254 <210> SEQ ID NO: 5
255 <211> LENGTH: 13
256 <212> TYPE: PRT
257 <213> ORGANISM: Artificial Sequence
259 <220> FEATURE:
260 <223> OTHER INFORMATION: Description of Artificial Sequence: derivatives
261     of segment of V3 domain in R2 strain
263 <220> FEATURE:
264 <221> NAME/KEY: VARIANT
265 <222> LOCATION: (3)..(12)
266 <223> OTHER INFORMATION: X = any natural or non-natural amino acid.
268 <400> SEQUENCE: 5
W--> 269 Pro Met Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gln
270   1           5           10
273 <210> SEQ ID NO: 6
274 <211> LENGTH: 30
275 <212> TYPE: PRT
276 <213> ORGANISM: Human immunodeficiency virus type 1
278 <220> FEATURE:
279 <223> OTHER INFORMATION: sequence of Phenetic 1 in V3 region
281 <400> SEQUENCE: 6
282 Asn Asn Thr Arg Lys Ser Ile His Ile Gly Pro Gly Arg Ala Phe Tyr
283   1           5           10           15
-----

```

RAW SEQUENCE LISTING

DATE: 09/16/2003

PATENT APPLICATION: US/09/762,261C

TIME: 16:45:03

Input Set : A:\h m jackson 5001 us.txt

Output Set: N:\CRF4\09162003\I762261C.raw

```

285 Thr Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys
286          20          25          30
289 <210> SEQ ID NO: 7
290 <211> LENGTH: 30
291 <212> TYPE: PRT
292 <213> ORGANISM: Human immunodeficiency virus type 1
294 <220> FEATURE:
295 <223> OTHER INFORMATION: sequence of Phenetic 2 in V3 region
297 <400> SEQUENCE: 7
298 Asn Asn Thr Arg Lys Ser Ile Ser Ile Gly Pro Gly Arg Ala Phe Tyr
299   1          5          10          15
301 Ala Thr Gly Glu Ile Ile Gly Asp Ile Arg Gln Ala His Cys
302          20          25          30
305 <210> SEQ ID NO: 8
306 <211> LENGTH: 30
307 <212> TYPE: PRT
308 <213> ORGANISM: Human immunodeficiency virus type 1
310 <220> FEATURE:
311 <223> OTHER INFORMATION: sequence of Phenetic 3 in V3 region
313 <400> SEQUENCE: 8
314 Asn Asn Thr Arg Lys Ser Ile Ser Ile Gly Pro Gly Arg Ala Phe Tyr
315   1          5          10          15
317 Ala Thr Gly Lys Ile Ile Gly Asp Ile Arg Gln Ala His Cys
318          20          25          30
321 <210> SEQ ID NO: 9
322 <211> LENGTH: 30
323 <212> TYPE: PRT
324 <213> ORGANISM: Human immunodeficiency virus type 1
326 <220> FEATURE:
327 <223> OTHER INFORMATION: sequence of Phenetic 4 in V3 region
329 <400> SEQUENCE: 9
330 Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro Gly Gln Ala Phe Tyr
331   1          5          10          15
333 Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys
334          20          25          30
337 <210> SEQ ID NO: 10
338 <211> LENGTH: 30
339 <212> TYPE: PRT
340 <213> ORGANISM: Human immunodeficiency virus type 1
342 <220> FEATURE:
343 <223> OTHER INFORMATION: sequence of Phenetic 5 in V3 region
345 <400> SEQUENCE: 10
346 Asn Asn Thr Arg Lys Ser Ile His Ile Gly Pro Gly Arg Ala Phe Tyr
347   1          5          10          15
349 Ala Thr Gly Lys Ile Ile Gly Asp Ile Arg Gln Ala His Cys
350          20          25          30
353 <210> SEQ ID NO: 11
354 <211> LENGTH: 30
355 <212> TYPE: PRT

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/762,261C

DATE: 09/16/2003
TIME: 16:45:04

Input Set : A:\h m jackson 5001 us.txt
Output Set: N:\CRF4\09162003\I762261C.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:5; Xaa Pos. 3,4,5,6,7,8,9,10,11,12